DOCUMENT RESUME

ED 326 910 CS 507 353

AUTHOR Stevens, Matthew D.; Hellweg, Susan A.
TITLE Toward a Standard of Communication Training Effectiveness Evaluation.
PUB DATE Nov 90
NOTE 35p.; Paper presented at the Annual Meeting of the Speech Communication Association (76th, Chicago, IL, November 1-4, 1990).
PUB TYPE Speeches/Conference Papers (150) -- Viewpoints (120)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Business Communication; *Evaluation Methods; Evaluation Needs; Institutional Cooperation; Speech Communication; Training; *Training Methods
IDENTIFIERS Applied Communication; Research Priorities; *Training Effectiveness

ABSTRACT

Communication training efforts in American business have increased steadily for the past several years. While this increase may be viewed as positive from several vantage points, it has not been matched by an increase in any systematic application of evaluation measures. Effective evaluation should take place at various levels. D. L. Kirkpatrick (1979) broke down evaluation efforts into the following four categories: reaction, learning, behavior, and results. From a survey of the literature, it is apparent that the most frequently used methods of evaluation currently are: (1) surveys to measure reaction and perceptions of results; (2) interviews to measure reaction, learning, knowledge of behavior, and perceptions of results; (3) tests to measure learning and behavior knowledge; (4) simulations and role plays to measure behavioral abilities; and (5) analysis of organization documents to assess results. To achieve acceptable standards of communication training evaluation, there must be more consolidated efforts between academicians and practitioners. Efforts directed toward developing and disseminating standardized guidelines of evaluation are greatly needed and should be a high priority for applied communication researchers. Further research specific to the communication training domain is needed to determine a time frame for evaluation so that researchers may be able to compare findings across organizational settings when identical instruments are used. Also, further testing of S. DeWine's Training Impact Questionnaire is needed to substantiate the instrument's validity. (Thirty-nine references are attached.) (MG)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED326910

Toward a Standard of Communication Training Effectiveness Evaluation

Matthew D. Stevens

and

Susan A. Hellweg

Department of Speech Communication

San Diego State University

San Diego, CA 92182-0300

(619) 594-5363/(619) 594-5610

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Matthew D. Stevens

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Applied Communication Section

Speech Communication Association Convention

Chicago, IL

November, 1990

CS507353

Running head: TRAINING EVALUATION

Abstract

Communication training efforts in American business have increased steadily for the past several years. While this increase may be viewed as positive from several vantage points, it has not been matched by an increase in any systematic application of evaluation measures. This essay explicates the domain of communication training, levels at which evaluation must take place, and the lack of and need for effective evaluations. Current evaluation practices employed by trainers and reported in the literature are examined, as well as prescriptions for the establishment of standards for the field. DeWine's Communication Training Impact Questionnaire is examined as a fundamental tool in the development of a comprehensive evaluation program.

Toward a Standard of Communication Training Effectiveness Evaluation

In recent years responsibility for communication training in organizations has been splintered and come under the auspices of several disciplines (Pace, Peterson, & Porter, 1984). In addition, it has been conducted by both external consultants and internal practitioners. For these reasons and others to be explored, there currently exists widespread disagreement about the effectiveness of communication training. Adding to this disparity is a paucity of tried and tested evaluation methods and measures. The purpose of this paper is to explore the lack of communication training evaluation standards and to point out directions for rectification. While several recent papers have called for increased evaluation efforts (Bell & Kerr, 1987; DeWine, 1987; Ford, 1989; Johnson & Kusmierek, 1987a and 1987b; Lashbrook, 1983), none has both thoroughly enumerated the reasons for the lack of fully developed measurement systems and forcefully called for researchers to rally around a common focus. This will be attempted here.

After communication training is defined and levels of training evaluation are described, the lack of evaluation efforts will be addressed. Evaluation needs

will be established and current efforts will be considered. Finally, issues of cooperation between academicians and practitioners in pursuit of evaluation standardization will be examined.

Communication Training Defined

Anastasi (1987) explained that communication training "Should not aim to make the participants aware or sensitive, or even knowledgeable, rather it should leave them with the ability to communicate better" (p. 2). The skills developed by communication trainers were identified by Ford (1989) as writing, oral communication, listening, interviewing, reading, conference management, group communication, nonverbal communication, dictating, sales presentations, supervisory communication, and assertive communication. Kelly (1986) explained that communication trainers are those who teach skills such as "public speaking, group facilitation, writing, and interpersonal communication" (p. 63).

Levels of Evaluation

Kirkpatrick (1979) first broke down evaluation efforts in a series of articles in 1959 and 1960 (reprinted in 1979), into four categories: reaction, learning, behavior, and results.

Reaction

The first level of evaluation is trainees' liking of training and is sometimes referred to as the "smile test" (Ford, 1989) or "happiness index" (Robinson & Robinson, 1989). Optimally, trainees' reactions to both content and trainer are measured, including instructor skills, training climate, course objectives, materials used, and organization. Kirkpatrick (1979) defended reaction evaluation, explaining that orders for future training are often based on one or two trainee's comments. He also noted that liking can predict attentiveness.

Learning

The second level of training evaluation concerns whether trainees remember (but not necessarily use) facts, principles, and behaviors taught. Two variables examined are knowledge gained and behavior learned (Ford, 1989). Knowledge gain may be measured with paper and pencil tests, but behavior learned requires more creative measurement such as role play and ability to explain behaviors to others.

Kirkpatrick (1979) explained that much learning evaluation is completed, but that it is the least relevant or practical level for an organization. Indeed, Burnaska (1976) found that while trainees

displayed knowledge gain after interpersonal communication skills training, no behavior change was identified. Kirkpatrick offered guidelines for learning assessment, recommending that each trainee be tested using pretest and posttest quantitative measures compared to a control group.

Behavior

Evaluation at the third level is complex and can be difficult to execute. Variables examined can be the desired outcomes of a training program. For example, if oral communication was the skill taught, improved ability to create and deliver speeches would be measured.

Kirkpatrick (1979) explained that behavior evaluation should meet the criteria for learning and should be done by one or more from among the trainee, the trainee's supervisor, subordinates, and peers. A criterion set up by Kirkpatrick but often ignored by practitioners is that assessment should be performed three or more months after training.

Results

Results are actual effects training has on an organization's functioning. Pace (1983) recommended that "Perceptions of supervisors, cost-effectiveness figures, [and] problems reported by supervisors" (p.

258) be observed. Ford (1989) listed such variables as trends in turnover, absences, sales, customer satisfaction indices, quantity and quality of performance, morale, cost reductions, grievance reductions, accident rates, amount of suggestions, employee satisfaction, and time use.

Alternative Levels

Several researchers since Kirkpatrick have developed their own nomenclature for levels of training evaluation, most renaming or only slightly altering the original categories. Pace (1983), for example, looked at trainee satisfaction (reaction), trainee information gain (learning), trainee skills acquisition (learning and behavior), and effect of skills on the organization (results). Lashbrook (1983) inverted levels one and two and renamed them "cognitive" and "affective," while leaving levels three and four unaltered.

The evaluation schema developed by Warr, Bird, and Rackham (1970) includes categories of context, input, reaction, and outcome, but is much broader than those previously discussed. Their fourth level, outcome, encompasses a facsimile of Kirkpatrick's four levels. The general failure of researchers to accept and advance Kirkpatrick's specific nomenclature illustrates the lack of standardization in evaluation. Since his

levels, if not his labels, are employed widely and mark the formation of the first comprehensive taxonomy, Kirkpatrick's labels will be used for the remainder of this paper, and references to levels will correspond to his categories in the order presented.

Evaluation Shortage

Several researchers claim there is a shortage of training evaluation in (and out) of the communication discipline (Denton, 1983; DeWine, 1987; Ford, 1989; Johnson & Kusmierek, 1987a and 1987b; Pullins, Gibbins, Wissinger, & DeWine, 1985; Ross & DeWine, 1987). In a survey of 106 human resource development training consultants who offer communication training, Johnson and Kusmierek (1987a and 1987b) found that consultants largely do not use experimental designs and seldom complete evaluations above level one. Looking at writing courses in 316 organizations, Denton (1983) found that only 9.4% attempt to measure effectiveness as it relates to results.

Smeltzer (1979) reported that less than 12% of 285 companies surveyed evaluate supervisory training programs. Bunker and Cohen (1977) asserted that evaluation is tremendously under researched and Ford (1989) found that overall, very few evaluation studies of communication skills training have been published.

When communication training is evaluated, the evaluation often covers only Kirkpatrick's levels one and two (DeWine, 1987). DeWine faulted researchers for ignoring the need for "A systematic assessment of behavioral changes...and the participants' ability to apply those changes" (p. 119).

An example comes from Blakeslee (1982), who administered a questionnaire six months after training (thus measuring long-term effects), but only asked subjects what changes in their organization's communication they perceived resulted from training. This subjective assessment falls short of answering DeWine's call. He supplemented this data with anecdotes from subjects. In this evaluation program, trainees' potential inability to see connections between training and their jobs was ignored. Coffman (1979) reported that her evaluation program consists of telling clients "Between now and our second session, each of you will gather information as to the results being produced by the particular training program being evaluated. You are to do this any way you think best" (p. 29). The ensuing session consists of clients reporting findings. This example illustrates the lack of systematic evaluation.

Trainers offer several reasons for their lack of

evaluation efforts. Bunker and Cohen (1978) discussed reasons given by in-house consultants, including unquestioned belief that training works, high costs of evaluation, a belief that better evaluations are not possible, and not wanting to risk failure. Two additional reasons may be that methodological expertise is often lacking and clients generally do not demand evaluation (Bell & Kerr, 1987).

Evaluation Needs

Rationale for conducting training evaluations is a need for evidence that training leads to greater organizational effectiveness. Kelley, Orgel, and Baer (1984) summarized Zenger and Hargis' four reasons for evaluating programs, saying:

> Training is relevant and functional only when it is evaluated; when budgets are trimmed, training departments that cannot substantiate their contribution to the company usually are...among the first to experience cutbacks; evaluations build credibility with departments otherwise reluctant to participate in training programs; and senior management wants to know the benefits as well as the costs of the training function. (p. 32)

Further benefits of training evaluation are increased

confidence in trainers' claims, increased legitimacy of the training function in organizations, justification for continued support and expenditures by clients, and valuable feedback for improving training methods.

Without concrete evidence from evaluation, advertisements such as the following example undermine the integrity of trainers: "Communication: The Key to Executive Performance," (from the August, 1978 issue of Training and Development Journal, p. 9).

For trainers to maintain legitimacy, the worth of their programs must be proven and documented. As trainers and consultants are not governed by a controlling body as are other professionals, it is important that trainers show their worth. As Bell and Kerr (1987) pointed out, evaluation practices are the key to continued support for training programs.

Pullins, et al. (1985) observed that since industry spends billions of dollars every year on communication training, the field needs effective means of evaluating the impact of such training. Estimates of U.S. expenditures on training run from $40 billion (DeWine, 1987) to $215 billion (Johnson & Kusmierek, 1987b) annually. With so much money being expended on training, it is essential that there be some proof of a return on this investment (Bell & Kerr, 1987; DeWine,

1987; Kelley, Orgel, & Baer, 1984; Kirkpatrick, 1979).

Finally, trainers cannot assume their programs are above need for modification. Arnold and McClure (1989) noted the importance of evaluation for improving communication training and outlined steps for using evaluation information to bolster programs. Rudolph and Johnson (1983) asserted the need to "appraise work in progress and when completed; correct performance; improve work being done and results secured" (p. 15). Improved training is further justified by the transformation of American society from an industrial to an information society (Naisbitt, 1984), where employee communication skills will be increasingly required.

The call for effective evaluation is not new. Goodacre (1957) called for it more than 30 years ago and his summons has echoed through the years from others (e.g., DeWine, 1987; Heisler, 1987; Johnson & Kusmierek, 1987a and 1987b; Kelley, Orgel, & Baer, 1984; Kirkpatrick, 1979; Lashbrook, 1983; Zenger & Hargis, 1982). Zenger and Hargis (1982) made the call, explaining that it is time to "take the plunge" for five reasons: interest in evaluation by managers and trainers is higher than ever, managers are now demanding return on investment (ROI) information more

than in the past, scientific methods are now mature, these methods can be counted on to produce objective results, and the current productivity crisis requires that training methods be improved.

Examination of Current Efforts

Despite the gloomy assessment offered thus far, there currently exist many methods of training evaluation. Through a review of these methods, however, the need for standardization becomes evident.

Evaluation Methods and Measures

From a survey of the literature, it is apparent that the most frequently used methods of evaluation currently are: surveys (questionnaires and checklists) to measure reaction and perceptions of results; interviews to measure reaction, learning, knowledge of behavior, and perceptions of results; tests to measure learning and behavior knowledge; simulations and role plays to measure behavioral abilities; and analysis of organization documents to assess results. An example of techniques used at different levels of evaluation comes from Bell and Kerr (1987). In evaluating their business communication training they used a questionnaire to measure level one, reaction, and a variety of pretests and posttests to measure level two, learning of writing, speaking, and dictation skills.

To measure level three behavior, Bell and Kerr (1987) used self-reports 3 to 23 months after training. Finally, the researchers claimed to measure level four, ROI, by looking at changes in work efficiency of trainees. Data collection was accomplished through self-report questionnaires. Where respondents reported improved relations with supervisors the researchers assumed increased efficiency.

Webb (1989) reported on her method of evaluating a public speaking training program. Whereas Bell and Kerr used only tests they created, Webb used a combination of previously published and newly created scales. The instruments Webb used were McCroskey's (1970) Personal Report of Communication Apprehension (PRCA); Hart, Carlson, and Eadie's (1980) Rhetorical Sensitivity Scale (Rhetsen Scale); Webb and Howay's (1987) Workshop Evaluation; and Webb's (1989) Public Speaking Quiz (PSQ), a 17-item test of speaking and listening competency questions. DeWine (1987) created the Training Impact Questionnaire to evaluate reaction and behavior. The instrument measures participants' perceptions of their ability to apply learned skills to their jobs.

A number of guidelines for the development of evaluation instrumentation are presented in the

literature. Phillips (1983) prescribed tests of validity and reliability, ease of administration, simplicity and brevity, and economy for evaluation instruments. Arnold and McClure (1989) identified six test designs for evaluating learning in the communication training context: immediate test, posttest only, pretest/posttest, time-series, Solomon Four-Group design, and multiple baseline design.

With an immediate test, the test is administered without delay after training and no control group is used. This lack of control group limits its usefulness. In posttest only and pretest/posttest designs, a control group is used, and is fairly well supported (e.g., Bunker & Cohen, 1978; Zenger & Hargis, 1982). The time-series design has the benefit of multiple tests. The Solomon Four-Group design is possibly the most valid of those discussed (Bunker & Cohen, 1978), but is not often used because of the many subjects and extra analysis required (Johnson & Kusmierek, 1987b). The multiple baseline design works well when a control group is difficult or impossible to obtain. While these tests are useful for measuring learning, behavior requires other evaluations.

At the behavioral level, Anastasi (1976) explained that communication "programs are best measured by the

skills they result in on the job" (pp. 31-31). This may be accomplished by taking Kirkpatrick's (1979) suggestion to have the trainee's supervisor, subordinates, and peers make assessments, a notion supported by Zenger and Hargis (1982). In her review of training evaluation articles, Ford (1989) found that in addition to the methods just discussed, trainers use interviews to both clarify questionnaire responses and collect stories and anecdotes about behavior.

Current evaluation efforts at the results level are not far beyond what they were 30 years ago when Kirkpatrick claimed that level four evaluation was beyond existing social scientific techniques. When level four evaluation is undertaken it generally consists of researchers looking at and interpreting sales figures, turnover and performance ratings (Ford, 1989), and time-use appraisals (Pace, 1983). Thus the analysis varies greatly by evaluator.

Finally, the amount of time allowed to pass between training and evaluation must be considered. Kirkpatrick (1979) recommended that level three be evaluated three or more months after training. DeWine (1987) prescribed that the Training IQ be administered two to four weeks after training. Anastasi (1976) recommended a two- to four-month delay. Blakeslee

(1982), using an ex-post-facto evaluation design, waited six months before measuring effects; Bell and Kerr (1987) completed evaluations between 3 and 23 months after training. These prescriptions suggest that there is not a fixed time frame for evaluation. This further limits the possibilities of comparing evaluation results from different training programs.

Before leaving this topic it should be noted that while many instructional designers and trainers develop and administer training and then develop an evaluation program, Robinson and Robinson (1989) propose something much different. They contend that training evaluation should be built into the training from the start and should occur at various stages of the training and after. They argue that no effective training can take place before the trainer knows what questions the learner must be able to answer in order to demonstrate that the training was successful.

Methods of Analysis

In the context of training assessments Kirkpatrick (1979) concluded that "To determine effectiveness, attempts must be made to measure in scientific and statistical terms" training outcomes (p. 89). However, he argued that in 1959 social science lacked the statistical sophistication necessary to quantify

training effectiveness. Kirkpatrick stated, as others have since (e.g., Arnold & McClure, 1989), that qualitative data does not lend itself to summarization and is not valuable for determining and documenting patterns of reaction. Some resolution to this dilemma may be at hand as Zenger and Hargis declared in 1982 that statistical techniques were sophisticated enough for use in evaluation.

Bell and Kerr (1987) contended that staff time is the most expensive resource an organization uses and that by documenting the amount of time spent on various tasks before and after training one can determine training effects. Their analysis consisted of simply administering pretests and posttests on skills taught and subsequently determining the percentage differences in time used for the tasks. Others use similar methods and employ t tests to measure changes. To quantify ROI using these methods one can apply formulas which translate time savings into dollar values, such as those proposed by Pace (1983).

To analyze reaction data researchers look at respondents' ratings using Likert-type scales. To analyze learning data, many researchers conduct t tests or calculate percentages of correct answers on tests and compare trainees' scores to a control group.

Behavioral data are commonly analyzed by calculating what percentage of respondents report using the new skill or what percentage of supervisors or peers report that the trainee uses the new skill. An alternate method of analysis is offered by Blakeslee (1982), who conducts content analysis of anecdotal data gathered through the use of a questionnaire.

Toward Better Evaluation Standards

Bridging Academia and Practice

To achieve acceptable standards of communication training evaluation, there must be more consolidated efforts between academicians and practitioners. In their analysis of academic programs for educating trainers, Pace, et al. (1984) found that training and development is taught in no less than eight different departments. These include behavioral science, business, communication, adult education, instructional media, education, counseling, and vocational education. Pace, et al. cautioned that academic programs need to work toward a shared focus instead of emphasizing the "less comprehensive and abstract models from specific academic disciplines" (p. 31).

As long as academic institutions maintain training education programs that operate independently of one another, there will continue to be errant assessments

of organizational training needs. Through collective efforts of schools of business and communication (to name one combination) more realistic appraisals of skills needed by trainers and the how-to of training evaluation could be made.

DeWine (1987) asserted that "The field of organizational communication needs to meet the demand for applications of organizational theory to the work world" (p. 113). Jarboe (1989a, 1989b) made a plea for communication departments to teach training as part of a communication consulting course.

While it would be naive to think that HRD training education could be brought under one academic roof, it is realistic to expect a melding of efforts in the future. The American Society for Training and Development proposed a theoretical focus for HRD training that could serve as a rallying point for training education (Pace, et al., 1984). Comeaux (1989) proposed that universities establish centers for applied communication research and consulting in hopes of bridging the gap between academia and business.

Communication trainers are grouped with HRD trainers in their professional organizations. The resulting problem from an evaluation perspective is that while some training lends itself to direct

evaluation, much communication training produces internal, not readily observable changes.

Another problem for practitioners is that training evaluation literature is scattered among professional journals in communication, personnel, management, training, and psychology. Even the most zealous training director could not be expected to read all applicable journals, thus recommendations for training evaluation and pilot studies of programs are read by only a fraction of practitioners.

Finally, within the training profession, trainers who might develop valuable evaluation programs are sometimes given cold receptions by top management. That is, a fundamental barrier to effective training evaluation is the failure of management to recognize the need for it (Bunker & Cohen, 1978; Jarboe, 1989b), or worse, a failure even to see a need for improved communication (Wasylik, Sussman, & Leri 1976).

As for professional associations, rather than create their own, communication trainers should strive for more influence within existing associations and should work toward self-governing divisions within them. With regard to training literature, in order to achieve consensus about evaluation within the training field, practitioners must secure an open and meaningful

dialogue. Rallying around a few common publications is the answer to this problem.

More cooperation between academicians and practitioners is also needed. Jarboe (1989b) pointed out that "Practitioners are considered 'soft' by researchers, and theorists are criticized by organizations for not being able to tackle the problems that real organizations face" (p. 23). Practitioners need to recognize the value of researcher's contributions to the training field and researchers need to conduct the kind of research that reports on exchanges in actual organizations (Goodall, 1984).

Some progress toward bridging the gap between academia and practice has been made. First, efforts such as the formulation of the Academy-Industry Task Force in Division IV of the International Communication Association have been noteworthy, the charge of this group being to sponsor convention panel programs to examine mutual concerns and to facilitate bridges between the two groups. Second, publications that are the result of joint efforts between academicians and practitioners seem to be appearing more frequently (e.g., Arnold & McClure, 1989; Bell & Kerr, 1987; Johnson & Kusmierek, 1987a & 1987b). Third, a recent survey of organizational communication textbooks (which

potentially could be used in consulting and training courses) indicated that students are being taught themes important in business communication, such as information exchange, conflict, and persuasion (Willer, Bell, & Andersen, 1987). As more universities add communication consulting courses, or at least emphasize consulting issues in existing organizational communication courses, current texts will help to equip future practitioners with needed skills. Finally, Ford (1989) called for researchers to more frequently publish information pertaining to their evaluations so that others might benefit.

Establishing a Standard

Blakeslee (1982) observed that the "day-to-day demands of the work environment make it almost impossible to do thoughtful evaluation," (p. 84) and noted the difficulties associated with validity and reliability in measurement design. However, he concluded that these obstacles can be overcome. Kirkpatrick (1979) asserted that "Whenever possible, it is suggested that training directors devise their own methods and techniques" for testing learning (level two). While trainers need to formulate or adapt specific questions that are specific to their training, there needs to be standard methods and techniques. In

support, Clegg (1987) found that nearly half of the large industrial corporations surveyed listed a lack of standards as the most pressing problem to in-house training program evaluation ventures.

Through the literature review here it becomes apparent that current evaluation programs do not address all four of Kirkpatrick's levels. There are, however, ample evaluation programs and guidelines that apply generalizable measures at two levels. One such schedule is DeWine's (1987) Training IQ, which measures reaction and behavior.

The Training IQ is an instrument designed primarily to assess trainees' "perceptions of their ability to apply skills learned in a training environment to their jobs" (DeWine, 1987, p. 118), belief that they have mastered the skill taught, and reactions to the training and trainer(s). It works toward accomplishing these tasks through analysis of reactions to 20 questions associated with two factors: "relationship of training to job," and "skilled performance." An overall reliability coefficient of .88 has been achieved for the measure, indicating that it is internally consistent. Predictive validity has been measured by correlating training evaluation and trainer credibility. Pullins, Gibbons, Wissinger, and

DeWine (1985) reported .73 correlation between the Training IQ and McCroskey's credibility scale. To provide additional validity for the instrument, DeWine studied 17 middle managers, comparing their Training IQ scores to assessments made by their supervisors. Moderate correlations of .57 for participant's perceptions of the application of skills and .58 for their evaluation of their mastery of skills were achieved. This report of criterion validity must be viewed cautiously as neither a control group nor a pretest was used and the sample was quite small.

Certain shortcomings of DeWine's measure should be noted. To use the measure is to assume that participants are capable of seeing the applicability of skills taught. This test group of middle managers would be expected to be at least moderately well educated and therefore could be expected to readily identify skill applicability, and there was no control for supervisor bias/liking in the survey.

The measure of perception of mastery, especially, needs to be viewed in light of the participant's level of performance prior to training. Also, the short lapse between training and testing (two to four weeks) may not be long enough, as several researchers (Anastasi, 1976; Bell & Kerr, 1987; Blakeslee, 1982;

Kirkpatrick, 1979) call for a much longer delay. Certain items of the Training IQ might not reasonably be expected to be measurable after only four weeks (e.g., "I didn't learn this skill in the training program, so I had to learn it on the job"). It should also be noted that the two factors extracted for the instrument accounted for only 35% of the variance.

Though the Training IQ has some weaknesses, it possesses several notable strengths. The reliability of the measure (.88), though needing further validation, is a strong point in its favor. DeWine tested the measure's internal validity on a sample of 267, giving credence to her reliability claim. The instrument measures participant's perception of skill applicability, so those who score high can be expected to find creative ways to apply the learned skill.

The inclusion of items which measure reaction (level one) as well as behavior (level three) add to the usefulness of the instrument. By grouping reaction measures with behavior measures (allowing time to pass after training), the trainer comes to conclusions based less on after-training euphoria than with other measures, thus overcoming criticism for evaluating participant's feelings immediately after training when spirits run high.

The DeWine (1987) instrument seems to be an effective tool for measuring behavior and reaction. Bolstered by qualitative interview data and further validity tests, it could be a valuable component of a comprehensive training evaluation program. Such a program must also include measures of learning and results (levels two and four).

To measure learning, a communication training evaluation program should include guidelines for testing. Specific questions must be tailored to the subject matter and organization, but a standardized format should be used (Kirkpatrick, 1979). To round out an evaluation agenda, results must be effectively appraised. Statistical analysis of quantitative organization outcomes should be used. Trainers should learn to conduct statistical tests of the data resulting from training programs to determine changes in organization outcomes or should seek expert assistance. And above all, systematization should be sought.

Summary and Conclusions

This essay has explored several facets of communication training evaluation efforts. Levels of evaluation were identified, the lack of and need for evaluation were discussed, and current efforts were

described. It became evident that a unified approach to evaluation is needed. Drawing on established practices and guidelines, a skeletal model of such a comprehensive approach was depicted. It was shown through this discussion that standardized guidelines of evaluation are attainable and desirable. Efforts directed toward developing and disseminating such standards are greatly needed and should be a high priority for applied communication researchers. Further research specific to the communication training domain is needed to determine a time frame for evaluation so as to be able to compare findings across organizational settings when identical instruments are used. Also, further testing of DeWine's Training IQ is needed to substantiate the instrument's validity.

References

Anastasi, T. E., Jr. (1976). Communications training. In R. L. Craig (Ed.), Training and development handbook: A guide to human resource development (pp. 31-1 - 31-32). New York: McGraw-Hill.

Anastasi, T. E., Jr. (1987). Communication training. In R. L. Craig (Ed.), Training and development handbook (pp. 740-755). New York: McGraw-Hill.

Arnold, W. E., & McClure, L. (1989). Communication training and development. New York: Harper & Row.

Bell, J. D., & Kerr, D. L. (1987). Measuring training results: Key to managerial commitment. Training and Development Journal, 41(1), 70-73.

Blakeslee, G. S. (1982). Evaluating a communications training program. Training and Development Journal, 36(11), 84-89.

Bunker, K. A., & Cohen, S. L. (1977). The rigors of training evaluation: A discussion and field demonstration. Personnel Psychology, 30, 525-541.

Bunker, K. A., & Cohen, S. L. (1978). Evaluating organizational training efforts: Is ignorance really bliss? Training and Development Journal, 32(8), 4-11.

Burnaska, R. F. (1976). The effects of behavior modeling training upon managers' behaviors and

employees' perceptions. Personnel Psychology, 29, 329-349.

Clegg, W. H. (1987). Management training evaluation: An update. Training and Development Journal, 41, 65-71.

Coffman, L. (1979). An easy way to effectively evaluate program results. Training and Development Journal, 33(8), 28-32.

Comeaux, P. (1989, November). Applied communication training: A model for integrating university teaching with communication consulting and research. Paper presented at the annual meeting of the Speech Communication Association, San Francisco, CA.

Denton, L. W. (1983). In-house training in business communication: A status report. In W. J. Buchholz (Ed.), Communication training and consulting in business, industry, and government (pp. 37-48). Urbana, IL: American Business Communication Association.

DeWine, S. (1987). Evaluation of organizational communication competency: The development of the Communication Training Impact Questionnaire. Journal of Applied Communication Research, 15, 113-127.

Ford, W. Z. (1989, November). Evaluating communication skills training in organizations. Paper presented at the annual meeting of the Speech Communication

Association, San Francisco, CA.

Goodacre, D. M. III. (1957). The experimental evaluation of management training: Principles and practice. Personnel, 3, 534-538.

Goodall, H. L. (1984). The status of communication studies in organizational contexts: One rhetorician's lament after a year-long odyssey. Communication Quarterly, 32, 133-147.

Heisler, W. J. (1987, August). Diagnosing managerial training needs: An exploratory look at a factor-analytic regression approach in the public sector. Paper presented at the annual meeting of the Academy of Management, San Francisco, CA.

Jarboe, S. (1989a, November). From practice to theory: Training the communication trainer. Paper presented at the annual meeting of the Speech Communication Association, San Francisco, CA.

Jarboe, S. (1989b). Teaching communication consulting and training (or, reminisces of a trainer). The Southern Communication Journal, 55, 22-41.

Johnson, J. R., & Kusmierek, L. A. (1987a, November). Evaluation research practices among human resource training consultants. Paper presented at the annual meeting of the Speech Communication Association, Boston, MA.

Johnson, J. R., & Kusmierek, L. A. (1987b). The status of evaluation research in communication training programs. Journal of Applied Communication Research, 15, 144-159.

Kelley, A. I., Orgel, R. F., & Baer, D. M. (1984). Evaluation: The bottom line is closer than you think. Training and Development Journal, 38(8), 32-37.

Kelly, J. W. (1986). Training in organizations: Career opportunities for communication. ACA Bulletin, 55, 63-67.

Kirkpatrick, D. L. (1979). Techniques for evaluating training programs. Training and Development Journal, 29(6), 78-92.

Lashbrook, V. J. (1983, November). An approach to designing training and development programs. Paper presented at the annual meeting of the Speech Communication Association, Washington, DC.

Naisbitt, J. (1984). Megatrends: Ten new directions transforming our lives. New York: Warner Books.

Pace, R. W. (1983). Organizational communication: Foundations for human resource development. Englewood Cliffs, NJ: Prentice Hall.

Pace, R. W., Peterson, B. D., & Porter, W. M. (1984, May). Taking the leadership: An analysis of human

resource training and development academic programs. Paper presented at the National Conference and Exposition of the American Society for Training and Development, Dallas, TX.

Phillips, J. J. (1983). Handbook of training evaluation and measurement methods. Houston, TX: Gulf Publishing.

Pullins, E., Gibbins, J., Wissinger, G., & DeWine, S. (1985, April). Organizational communication training and credibility: Evaluation of impact. Paper presented at the annual meeting of the Central States Communication Association, Indianapolis, IN.

Robinson, D. G., & Robinson, J. C. (1989). Training for impact. San Francisco: Jossey Bass.

Ross, M. B., & DeWine, S. (1987, May). Organizational communication training and trainer credibility: Impact on learning and job applications. Paper presented at the annual meeting of the International Communication Association, Montreal, Quebec, Canada.

Rudolph, E. E. & Johnson, B. R. (1983). Communication consulting: Another teaching option. Urbana, IL: ERIC Clearinghouse.

Smeltzer, L. R. (1979). Do you really evaluate, or just talk about it? Training, 16(8), 6-8.

Warr, P, Bird, M., & Rackham, N. (1970). Evaluation of

management training. London: Gower Press.

Wasylik, J. E., Sussman, L., & Leri, R. P. (1976). Communication training as perceived by training personnel. Communication Quarterly, 24, 32-38.

Webb, L. (1989). A program of public speaking training: One consultant's approach. The Southern Communication Journal, 55, 72-86.

Willer, L. R., Bell, K. D., & Andersen, P. A. (1987). Is what we teach about organizational communication what they practice in organizations? Journal of Applied Communication Research, 15, 95-112.

Zenger, J. H., & Hargis, C. (1982). Assessing training results: It's time to take the plunge. Training and Development Journal, 36(1), 11-16.